Multi-rate Sub-Nyquist Spectrum Sensing in Cognitive Radios

Hongjian Sun, *Member, IEEE*, A. Nallanathan*, *Senior Member, IEEE*, Jing Jiang, *Member, IEEE*, and Cheng-Xiang Wang, *Senior Member, IEEE*

arXiv:1302.1489v1 [cs.IT] 6 Feb 2013

Abstract

Wideband spectrum sensing is becoming increasingly important to cognitive radio (CR) systems for exploiting spectral opportunities. This paper introduces a novel multi-rate sub-Nyquist spectrum sensing (MS^3) system that implements cooperative wideband spectrum sensing in a CR network. MS^3 can detect the wideband spectrum using partial measurements without reconstructing the full frequency spectrum. Sub-Nyquist sampling rates are adopted in sampling channels for wrapping the frequency spectrum onto itself. This significantly reduces sensing requirements of CR. The effects of sub-Nyquist sampling are considered, and the performance of multi-channel sub-Nyquist samplings is analyzed. To improve its detection performance, sub-Nyquist sampling rates are chosen to be different such that the numbers of samples are consecutive prime numbers. Furthermore, when the received signals at CRs are faded or shadowed, the performance of MS^3 is analytically evaluated. Numerical results show that the proposed system can significantly enhance the wideband spectrum sensing performance while requiring low computational and implementation complexities.

Index Terms

Cognitive radio, Spectrum sensing, Sub-Nyquist sampling, Rayleigh distribution, Log-normal distribution.

H. Sun and A. Nallanathan* are with the Department of Electronic Engineering, King's College London, London, WC2R 2LS, UK. (Email: hongjian.sun@kcl.ac.uk; nallanathan@ieee.org)

J. Jiang is with Center for Communication Systems Research, University of Surrey, Guildford, GU2 7XH, UK. (Email: jing.jiang@surrey.ac.uk)

C.-X. Wang is with Joint Research Institute for Signal and Image Processing, School of Engineering & Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK. (Email: Cheng-Xiang.Wang@hw.ac.uk)

I. INTRODUCTION

The radio frequency (RF) spectrum is a scarce natural resource, currently regulated by government agencies. Under the current policy, the primary user (PU) of a particular spectral band has exclusive rights to the licensed spectrum. With the proliferation of wireless services, the demands for the RF spectrum are continually increasing. On the other hand, it has been reported that localized temporal and geographic spectrum utilization efficiency is extremely low. For example, it has been reported that the maximal occupancy of the spectrum between 30 MHz and 3 GHz is only 13.1% and its average occupancy is 5.2% in New York City [1]. The spectral under-utilization can be addressed by allowing secondary users to access a licensed band when the PU is absent. Cognitive radio (CR) has become one promising solution for realizing this goal [2], [3].

A crucial requirement of CRs is that they must rapidly fill spectral holes without causing harmful interference to the PUs. This ability is dependent upon spectrum sensing, which is considered as one of the most critical components in a CR system. In a multipath or shadow fading environment, the signal-to-noise ratio (SNR) of the primary signal as received at CRs can be severely degraded, which will not only lead to unreliable spectrum sensing results, but will also reduce the capacity of the CR network due to the decreased data transmission time per frame. In such a scenario, cooperative spectrum sensing could increase the reliability of spectrum sensing by exploiting spatial diversity. In our previous work [4], [5], centralized cooperative spectrum sensing frameworks have been developed for improving the reliability of spectrum sensing. However, these studies only considered narrowband spectrum sensing techniques, the extension to wideband cooperative spectrum sensing requires yet a different approach.

To exploit more spectral opportunities over a large range of frequencies (e.g., 10 kHz $\sim$ 10 GHz), a CR system needs some essential components, i.e., wideband antenna, wideband RF front end, and high speed analog-to-digital converter (ADC). Yoon *et al.* [6] have shown that the -10 dB bandwidth of the newly designed antenna can be 14.2 GHz. Hao and Hong [7] designed a compact highly selective wideband bandpass filter with a bandwidth of 13.2 GHz. **In [8], Bevilacqua and Niknejad designed a wideband CMOS low-noise amplifier with**

the bandwidth of approximately 10 GHz. In contrast, the development of ADC technology is relatively behind. To the best of our knowledge, when we require an ADC to have a high resolution and a reasonable power consumption, the achievable sampling rate of the current ADC is 3.6 Gsps [9]. Obviously, ADC becomes a bottleneck in such a wideband system. Even if there exists ADC with more than 20 Gsps sampling rate, the real-time digital signal processing of 20 Gb/s of data could be very expensive.

In previous work, Quan *et al.* [10], [11] proposed a multiband joint detection (MJD) approach that can sense the primary signal over a wide frequency range. It has been shown that MJD has superior performance for multiband spectrum sensing. In [12], Tian and Giannakis studied a wavelet detection approach, which could adapt parameters to a dynamic wideband spectrum. Furthermore, they cleverly introduced compressed sensing (CS) theory to implement wideband spectrum sensing by using sub-Nyquist sampling techniques in the classic paper [13]. Later on, the CS-based approach has attracted many talented-researchers' attention in [14]–[24] owning to its advantage of using fewer samples closer to the information rate, rather than the inverse of the bandwidth, to perform wideband spectrum sensing. In [14], Tian *et al.* studied cyclic spectrum sensing techniques with high robustness against sampling rate reduction and noise uncertainty. In [15], Zeng *et al.* proposed a distributed CS-based spectrum sensing approach for cooperative multihop CR networks. In our previous work [25], [26], to save system energy, adaptive CS-based spectrum sensing approaches were proposed that could find the best spectral recovery with high confidence. Unfortunately, using CS-based approaches, the spectral recovery may cause high computational complexity, leading to a high spectrum sensing overhead due to the restricted computational resources in CRs.

In this paper, we introduce a multi-rate sub-Nyquist spectrum sensing (MS^3) approach for cooperative wideband spectrum sensing in a CR network. Because the spectral occupancy is low, sub-Nyquist sampling is induced in each sampling channel to wrap the sparse spectrum occupancy map onto itself. The sensing requirements are therefore significantly reduced. We then analyze the effects caused by sub-Nyquist sampling, and represent the test statistic using a reduced data set obtained from multi-channel sub-Nyquist sampling. Furthermore, we propose to

use different sampling rates in different sampling channels for improving the spectrum sensing performance. Specifically, in the same observation time, the number of samples in multiple sampling channels are chosen as different consecutive prime numbers. In addition, the performance of MS^3 for combining faded or shadowed signals is analyzed, and the closed-form bounds for the average probabilities of false alarm and detection are derived. The key advantage of MS^3 is that the wideband spectrum can be detected directly from a few sub-Nyquist samples without spectral recovery. Compared to the existing spectrum sensing methods, MS^3 can achieve better wideband spectrum sensing performance with a relatively lower implementation complexity.

The rest of the paper is organized as follows. Section II introduces the signal model. In Section III, we propose the wideband spectrum sensing approach, i.e., MS^3. The performance analysis of MS^3 for combining faded signals is given in Section IV. Section V presents simulation results, and conclusions are given in Section VI.

II. Preliminary

Consider that all CRs keep quiet during the spectrum sensing interval as enforced by protocols, e.g., at the medium access control (MAC) layer [10]. Therefore, the observed spectral energy arises only from PUs and background noise. The bandwidth of the signal as received at CRs is W (Hertz). Over an observation time T, if the sampling rate f ($f \geq 2W$) is adopted to sample the received signal, a sequence of Nyquist samples will be obtained with the length of $JN \triangleq fT$. This sequence is then divided into J equal-length segments where N denotes the number of Nyquist samples per segment (both J and N are chosen to be natural numbers). If we use $x_{\mathrm{c},i}(t)$ ($t \in [0, T]$) to represent the continuous-time signal received at CR i, after Nyquist sampling, the sampled signal can be denoted by $x_i[n] = x_{\mathrm{c},i}(n/f),\ n = 0, 1, \cdots, JN-1$. At CR i, the sampled signal of segment j ($j \in [1, J]$) can be written as

$$x_{i,j}[n] = \begin{cases} x_{\mathrm{c},i}(n/f), & n = (j-1)N, (j-1)N+1, \cdots, jN-1 \\ 0, & \text{Otherwise}. \end{cases} \tag{1}$$

The discrete Fourier transform (DFT) spectrum of the sampled signal of segment j is given by

$$X_{i,j}[k] = \sum_{n=0}^{N-1} x_{i,j}[n] e^{-\jmath 2\pi kn/N}, \ k = 0, 1, \cdots, N-1 \tag{2}$$

where $\jmath = \sqrt{-1}$. We model spectrum sensing on a frequency bin k as a binary hypothesis test, i.e., $\mathcal{H}_{0,k}$ (absence of PUs) and $\mathcal{H}_{1,k}$ (presence of PUs) [11]:

$$X_{i,j}[k] = \begin{cases} Z_{i,j}[k], & \mathcal{H}_{0,k} \\ H_{i,j}[k]S_{i,j}[k] + Z_{i,j}[k], & \mathcal{H}_{1,k} \text{ or } k \in \Omega_i \end{cases} \tag{3}$$

where $Z_{i,j}[k]$ is complex additive white Gaussian noise (AWGN) with zero mean and variance $\delta_{i,k}^2$, i.e., $Z_{i,j}[k] \sim \mathcal{CN}(0, \delta_{i,k}^2)$, $H_{i,j}[k]$ denotes the discrete frequency response between the PU and CR i, $S_{i,j}[k]$ is assumed to be a deterministic signal sent by the PU on the frequency bin k, and Ω_i denotes the spectral support such that $\Omega_i = \{k|\text{PU presents at } X_{i,j}[k]\}$. For simplicity, in the rest of the paper, we assume that the noise variance of the DFT spectrum is normalized to be 1. The observation time T is chosen to be smaller than the channel coherence time so that the magnitude of $H_{i,j}[k]$ remains constant within T for one CR, i.e., constant $|H_{i,j}[k]|$ regarding the segment number j.

Because an energy detector does not require any prior information about the transmitted primary signal while having lower complexity than other spectrum sensing approaches [27], we consider the energy detection approach in this paper. The received signal energy can be calculated as

$$E_i[k] = \sum_{j=1}^{J} |X_{i,j}[k]|^2, \quad k = 0, 1, \cdots, N-1. \tag{4}$$

The decision rule for energy detection approach is then given by

$$E_i[k] \underset{\mathcal{H}_{0,k}}{\overset{\mathcal{H}_{1,k}}{\gtrless}} \lambda_k, \quad k = 0, 1, \cdots, N-1 \tag{5}$$

where λ_k is the detection threshold for the frequency bin k. Here, it is noteworthy to emphasize that, after Fourier transform, the energy detection is done on each frequency bin in the frequency domain. Thus, the noise in high frequencies should not affect the energy detection in low frequencies, and vice versa. The benefit of frequency-domain energy detection is that the

detection performance depends on the SNR *on a single frequency bin, regardless of the noise in the other frequencies* (e.g., high frequency noise due to wideband sensing). To be specific, the signal energy on frequency bin k can be modeled by [27]

$$E_i[k] \sim \begin{cases} \chi^2_{2J}, & \mathcal{H}_{0,k} \\ \chi^2_{2J}(2\gamma_i[k]), & \mathcal{H}_{1,k} \end{cases} \tag{6}$$

where $\gamma_i[k] \triangleq \frac{\mathbb{E}\left(|H_i[k]S_i[k]|^2\right)}{\delta^2_{i,k}}$ denotes the SNR on the frequency bin k at CR i, χ^2_{2J} denotes central chi-square distribution, and $\chi^2_{2J}(2\gamma_i[k])$ denotes non-central chi-square distribution. Both of these distributions have $2J$ degrees of freedom and $2\gamma_i[k]$ denotes a non-centrality parameter. Here, the noise has variance $\delta^2_{i,k}$, measured bandwidth $\frac{f}{N}$, and noise temperature $T_n = \frac{\delta^2_{i,k}N}{fK_B}$ where K_B denotes the Boltzmann constant. The probabilities of false alarm and detection are given by [27]

$$P_{\mathrm{f},i,k} = \Pr(E_i[k] > \lambda_k | \mathcal{H}_{0,k}) = \frac{\Gamma(J, \frac{\lambda_k}{2})}{\Gamma(J)} \tag{7}$$

$$P_{\mathrm{d},i,k} = \Pr(E_i[k] > \lambda_k | \mathcal{H}_{1,k}) = Q_J\left(\sqrt{2\gamma_i[k]}, \sqrt{\lambda_k}\right) \tag{8}$$

where $\Gamma(a)$ denotes the gamma function, $\Gamma(a,x)$ denotes the upper incomplete gamma function, and $Q_u(a,x)$ is the generalized Marcum Q-function defined by $Q_u(a,x) = \frac{1}{a^{u-1}} \int_x^\infty t^u e^{-\frac{a^2+t^2}{2}} I_{u-1}(at)dt$ in which $I_v(a)$ is the v-th order modified Bessel function of the first kind.

III. MULTI-RATE SUB-NYQUIST SPECTRUM SENSING

It is difficult to realize wideband spectrum sensing, because it requires a high speed ADC for Nyquist rate sampling. We will now present an MS^3 system using multiple low-rate samplers to implement wideband spectrum sensing in a CR network.

A. System Description

Consider that there are v synchronized CRs collaborating for wideband spectrum sensing, and the fusion center (FC) is one of the CRs which has either greater computational resources or longer battery life than other CRs. Due to low spectral occupancy [13], the received signals at CRs are often sparse in the frequency domain. Here, we assume that the Nyquist DFT spectrum, $\overrightarrow{X_{i,j}} \in \mathbb{C}^N$, is s-sparse ($s \ll N$), which means that only the largest s out of N components cannot

be ignored. The spectral sparsity level, i.e., s, can be obtained from either sparsity estimation [16] or system initialization (e.g., by long term spectral measurements). As shown in Fig. 1, MS^3 consists of several CRs, each of which has one wideband filter, one low-rate sampler, and a fast Fourier transform (FFT) device. The wideband filters are set to have bandwidth of W. MS^3 can be described as follows:

1) The FC allocates different sub-Nyquist sampling rates to different CRs.
2) CRs perform sub-Nyquist samplings in the observation time T.
3) The sub-Nyquist DFT spectrum is calculated by using sub-Nyquist samples and FFT device[1].
4) The signal energy vectors are formed by using the sub-Nyquist DFT spectrum.
5) The CRs transmit these signal energy vectors to the FC by using a dedicated common control channel in a band licensed to the CR network [28].
6) The received data from all CRs is fused in the FC to form a test statistic.
7) The FC chooses the detection threshold and performs binary hypothesis tests.
8) The FC shares the detection results with all CRs.

B. Sub-Nyquist Sampling and Data Combining

At CR i, we use sub-Nyquist rate f_i ($f_i < 2W \leq f$) to sample the continuous-time signal $x_{\mathrm{c},i}(t)$. The sampled signal can be denoted by $y_i[n] = x_{\mathrm{c},i}(n/f_i),\ n = 0, 1, \cdots, JM_i - 1$ where $JM_i = f_i T$ and M_i is assumed to be a natural number. The sampled signal is then divided into J equal-length segments. The segment j ($j \in [1, J]$) can be written as

$$y_{i,j}[n] = \begin{cases} x_{\mathrm{c},i}(n/f_i), & n = (j-1)M_i, (j-1)M_i + 1, \cdots, jM_i - 1 \\ 0, & \text{Otherwise}. \end{cases} \tag{9}$$

The DFT spectrum of the sampled signal of segment j ($j \in [1, J]$) can be given by

$$Y_{i,j}[m] = \sum_{n=0}^{M_i - 1} y_{i,j}[n] e^{-\jmath 2\pi mn/M_i},\ m = 0, 1, \cdots, M_i - 1 \tag{10}$$

[1]Jointly considering wideband spectrum sensing and spectrum reuse in CRs, we use FFT devices for distinguishing different frequencies in order to reuse some un-occupied frequencies. Here, the use of FFT will cause additional complexity of $\mathcal{O}(M \log M)$ and memory storage increment, if M denotes the number of FFT points.

With the aid of Poisson summation formula [29], the DFT spectrum of sub-Nyquist samples can be represented by the DFT spectrum of Nyquist samples (as proved in Appendix A):

$$Y_{i,j}[m] = \frac{M_i}{N} \sum_{l=-\infty}^{\infty} X_{i,j}[m + lM_i], \quad m = 0, 1, \cdots, M_i - 1. \tag{11}$$

According to (3) and (11), the spectral support of the sub-Nyquist spectrum $\overrightarrow{Y_{i,j}}$ can be given by

$$\Omega_{\mathrm{s},i} = \{m | m = |k|_{\ \mathrm{mod}\ (M_i)}, \ \ k \in \Omega_i\}. \tag{12}$$

One risk caused by sub-Nyquist sampling is the signal overlap in $Y_{i,j}[m]$. However, when we choose parameters in $JN = fT$ such that $N \gg s$ and let the sub-Nyquist sampling rate satisfy $M_i \sim \mathcal{O}(\sqrt{N})$, the probability of signal overlap is very small (as proved in Appendix B). In such a scenario, we concentrate on considering two cases: no signal on m and one signal on m. In the latter case, only a single l is active in (11), and the other terms in the summation of (11) can be modeled as noise by using (3). Thus, the following equation holds from (3) and (11):

$$Y_{i,j}[m] = \frac{M_i}{N} X_{i,j}[m + lM_i] + \frac{M_i}{N} \sum_{\nu \neq l} Z_{i,j}[m + \nu M_i], \quad m + lM_i \in \Omega_i \tag{13}$$

where l is an unknown integer within $[0, N/M_i - 1]$. Furthermore, using (3) and (13), we can model the DFT spectrum of sub-Nyquist samples by

$$\sqrt{\frac{N}{M_i}} Y_{i,j}\left[|k|_{\ \mathrm{mod}\ (M_i)}\right] \sim \begin{cases} \mathcal{CN}\left(0, \delta^2_{\mathrm{s},i,k}\right), & k \notin \Omega_i \\ \mathcal{CN}\left(\sqrt{\frac{M_i}{N}} H_{i,j}[k] S_{i,j}[k], \delta^2_{\mathrm{s},i,k}\right), & k \in \Omega_i \end{cases} \tag{14}$$

where $\delta^2_{\mathrm{s},i,k}$ is the noise variance of sub-Nyquist DFT spectrum, and can be given by using (11)

$$\delta^2_{\mathrm{s},i,k} = \underbrace{\left\lceil \frac{N}{M_i} \right\rceil}_{No.\ of\ sums} \left(\frac{M_i}{N} \underbrace{\sqrt{\frac{N}{M_i}}}_{Scaling\ of\ Y_{i,j}} \right)^2 \delta^2_{i,k} \approx \delta^2_{i,k} \tag{15}$$

where $\lceil \frac{N}{M_i} \rceil$ (the smallest integer not less than $\frac{N}{M_i}$) denotes the number of summations in (11).

The signal energy of sub-Nyquist DFT spectrum in each CR node is then calculated by

$$E_{\mathrm{s},i}[m] = \sum_{j=1}^{J} |Y_{i,j}[m]|^2, \quad m = 0, 1, \cdots, M_i - 1. \tag{16}$$

which can be modeled by using (14) and (16) as

$$\frac{N}{M_i}E_{\mathrm{s},i}\left[|k|_{\bmod(M_i)}\right] \sim \begin{cases} \chi^2_{2J}, & k \notin \Omega_i \\ \chi^2_{2J}\left(2\frac{M_i}{N}\gamma_i[k]\right), & k \in \Omega_i. \end{cases} \tag{17}$$

We note that, due to the sub-Nyquist sampling, the noise will be folded from the whole bandwidth onto all signals of interest as shown in (13). As a result, comparing (17) with (6), we find that the received SNR in the sub-Nyquist sampling channel i will degrade from γ_i to $\frac{M_i}{N}\gamma_i$. This SNR degradation depends on the ratio between the number of samples at the sub-Nyquist rate and the number of samples at the Nyquist rate (i.e., $\frac{M_i}{N}$).

In MS^3, the signal energy vectors at CRs will then be collected at the FC. Finally, we form a test statistic by

$$\widehat{E_{\mathrm{s}}}[k] = \sum_{i=1}^{v}\frac{N}{M_i}E_{\mathrm{s},i}[|k|_{\bmod(M_i)}], \quad k = 0, 1, \cdots, N-1. \tag{18}$$

In Fig. 2, we give an illustration of the above test statistic for a practical ASTC DTV signal. To test whether the PU is present or not, we adopt the following decision rule:

$$\widehat{E_{\mathrm{s}}}[k] \underset{\mathcal{H}_{0,k}}{\overset{\mathcal{H}_{1,k}}{\gtrless}} \lambda_k, \quad k = 0, 1, \cdots, N-1. \tag{19}$$

Let $\Omega_{\mathrm{A},i}$ denote a set of aliased frequencies (i.e., false frequencies appear as mirror images of the original frequencies around the sub-Nyquist sampling frequency), and $\Omega_{\mathrm{U},i}$ represent a set of unaffected/unoccupied frequencies:

$$\Omega_{\mathrm{A},i} \triangleq \left\{k \middle| m = |k|_{\bmod(M_i)}, \quad m \in \Omega_{\mathrm{s},i}, k \notin \Omega_i\right\} \tag{20}$$

$$\Omega_{\mathrm{U},i} \triangleq \left\{k \middle| m = |k|_{\bmod(M_i)}, \quad m \notin \Omega_{\mathrm{s},i}, k \notin \Omega_i\right\}. \tag{21}$$

Using (17), we can model the test statistic of (18) as

$$\widehat{E_{\mathrm{s}}}[k] \sim \begin{cases} \chi^2_{2Jv}, & k \in \Omega_{\mathrm{U}} \\ \chi^2_{2Jv}\left(\frac{2}{N}\sum\limits_{i\in\Upsilon}^{|\Upsilon|=p} M_i\gamma_i[k]\right), & k \in \Omega_{\mathrm{A}} \\ \chi^2_{2Jv}\left(\frac{2}{N}\sum\limits_{i=1}^{v} M_i\gamma_i[k]\right), & k \in \Omega \end{cases} \tag{22}$$

where $\Omega_{\mathrm{U}} \triangleq \cap_{i=1}^{v}\Omega_{\mathrm{U},i}$, $\Omega_{\mathrm{A}} \triangleq \cup_{i=1}^{v}\Omega_{\mathrm{A},i}$, $\Omega \triangleq \cap_{i=1}^{v}\Omega_i$, $\Upsilon \triangleq \{i | m = |k|_{\mod(M_i)}, m \in \Omega_{\mathrm{s},i}, k \notin \Omega_i\}$ denotes the set of CRs who have aliased frequency on the frequency bin k, and $|\Upsilon| = p$ denotes the cardinality of the set Υ (equivalently the number of CRs that have aliased frequencies on the frequency bin k).

In (22), $k \in \Omega_{\mathrm{U}}$ and $k \in \Omega_{\mathrm{A}}$ are two extreme cases under the hypothesis $\mathcal{H}_{0,k}$. The former case denotes there is no aliased frequency on the frequency bin k, while the latter case represents there are maximum number of aliased frequencies (i.e., p) on the frequency bin k. Thus, the former one is the best case while the latter one is the worst case for signal detection under the hypothesis $\mathcal{H}_{0,k}$. The probability of false alarm on the frequency bin k can therefore be bounded by using (22)

$$\frac{\Gamma(Jv, \frac{\lambda_k}{2})}{\Gamma(Jv)} \leq P_{\mathrm{f},k} \leq Q_{Jv}\left(\sqrt{\frac{2}{N}\sum_{i\in\Upsilon}^{|\Upsilon|=p} M_i\gamma_i[k]}, \sqrt{\lambda_k}\right). \tag{23}$$

We note that the problem of minimizing the probability of false alarm can be transformed to minimize the parameter p, which depends on several factors, e.g., the sampling rates of CRs. Using the same sub-Nyquist sampling rates in MS[3] is not recommended as it could lead to $p = v$, resulting in the maximum of the probability of false alarm. As we will see in the following subsection, the parameter p can be minimized by using different sampling rates at CRs.

C. Multi-rate Sub-Nyquist Spectrum Sensing

To improve the detection performance of sub-Nyquist sampling system in the preceding subsection, we should analyze the influence of sampling rates. Firstly, we consider the case of spectral sparsity level $s = 1$, which means that only one frequency bin $k_1 \in \Omega$ is occupied by the PU.

Lemma 1: If the numbers of samples in multiple CRs, i.e., $M_1, M_2, ..., M_v$, are different primes, and meet the requirement of

$$M_iM_j > N, \quad \forall\, i \neq j \in [1, v] \tag{24}$$

then two or more CRs cannot have mirrored frequencies in the same frequency bin.

The proof of Lemma 1 is given in Appendix C.

Secondly, considering the spectral sparsity level $s \geq 2$, we find that, if the conditions in Lemma 1 are satisfied, the parameter p in (22) is bounded by s. It is because only one CR can map the original frequency bin $k_j \in \Omega_i$ to the aliased frequency in Ω_{A}, and the cardinality of the spectral support Ω_i is s. Therefore, we obtain the detection performance of MS^3 as Theorem 1.

Theorem 1: In MS^3, if the numbers of samples in multiple CRs, i.e., $M_1, M_2, \cdots, M_v$, are different consecutive primes, and meet the requirement of $M_i M_j > N,\ \forall\ i \neq j \in [1, v]$, using the decision rule of (19) the probabilities of false alarm and detection have the following bounds:

$$\frac{\Gamma(Jv, \frac{\lambda_k}{2})}{\Gamma(Jv)} \leq P_{\mathrm{f},k} \leq Q_{Jv}\left(\sqrt{\frac{2}{N}\sum_{i\in\Upsilon}^{|\Upsilon|=s} M_i\gamma_i[k]}, \sqrt{\lambda_k}\right) \tag{25}$$

$$P_{\mathrm{d},k} \geq Q_{Jv}\left(\sqrt{\frac{2}{N}\sum_{i=1}^{v} M_i\gamma_i[k]}, \sqrt{\lambda_k}\right). \tag{26}$$

Proof: Using (23) and the bound $|\Upsilon| = p \leq s$, (25) follows. Furthermore, when the energy of one spectral component in Ω maps to another spectral component in Ω, the probability of detection will increase. Thus, the inequality of (26) holds. □

Remark 1: It can be seen from Theorem 1 that the sampling rates in MS^3 can be much lower than the Nyquist rate because of $M_i \sim \mathcal{O}(\sqrt{N})$. By (25) we note that the probability of false alarm increases when the spectral sparsity s increases. In addition, the higher average sampling rate will lead to better detection performance. This is because the probability of signal overlap in the aliased spectrum can be reduced with a larger M_i in each sampling channel as our discussions in Section III-B. By (25) and (26), we can see that using more sampling channels (i.e., v), the detection performance can be improved. It should be emphasized that there is no closed-form expression for the probabilities in Theorem 1. This is because the number of CRs that have aliased frequencies on the frequency bin k cannot be predicted. Moreover, we note that the upper and lower bounds in Theorem 1 can be easily computed because the Marcum-Q function can be efficiently computed using power series expansions [30]. Under the Neyman-Pearson

criterion, we should design a test with the constraint of $P_{\mathrm{f},k} \leq \alpha$. In such a scenario, we must let the upper bound of (25) to be α and solve the detection threshold λ_k from the inverse of the Marcum-Q function. It has been shown in [31] that the detection threshold can be calculated with low computational complexity. In addition, to calculate the detection threshold, the noise power is required to be known at the FC.

IV. Combination of Faded Signals

As shown in Section III, the combining procedure in the FC is to sum up all non-faded signals at CRs and make final decisions. In this section, we investigate the combination of faded signals at CRs using the same approach as shown in (18). We assume that the received primary signals at different CRs are independent and identically distributed (i.i.d.), and are faded subject to either Rayleigh or log-normal distribution.

Solving the distribution of the sum of weighted independent random variables in (18) is not trivial. Hence, we use the sum of uniformly weighted random variables to approximate the sum of different weighted random variables in Theorem 1:

$$\frac{2\sum_{i=1}^{v} M_i\gamma_i}{N} \simeq \frac{2\overline{M}}{N}\sum_{i=1}^{v}\gamma_i = \psi\gamma_{\mathrm{v}}, \qquad \frac{2\sum_{i\in\Upsilon}^{|\Upsilon|=s} M_i\gamma_i}{N} \simeq \frac{2\overline{M}}{N}\sum_{i\in\Upsilon}^{|\Upsilon|=s}\gamma_i = \psi\gamma_{\mathrm{s}} \tag{27}$$

where $\overline{M}$ is the average M_i over multiple CRs, $\psi \triangleq \frac{2\overline{M}}{N}$, $\gamma_{\mathrm{v}} \triangleq \sum_{i=1}^{v}\gamma_i$, and $\gamma_{\mathrm{s}} \triangleq \sum_{i\in\Upsilon}^{|\Upsilon|=s}\gamma_i$. We note that the above approximation accuracy mainly depends on $\frac{|\overline{M}-M_i|}{N}$, where smaller $\frac{|\overline{M}-M_i|}{N}$ corresponds to more accurate approximation. Since $M_1, M_2, ..., M_v$ are chosen to be v different consecutive prime numbers and the distance between primes could be very small compared to N, the parameter $\frac{|\overline{M}-M_i|}{N}$ will approach to zero as N increases. Thus, the above approximation has little impact on the final result.

A. Rayleigh Distribution

If the magnitudes of received signals at different CRs follow Rayleigh distribution, then the SNRs will follow exponential distribution. Hence, γ_{v} and γ_{s} follow Gamma distributions:

$$f(\gamma_{\mathrm{v}}) = \frac{\gamma_{\mathrm{v}}^{v-1}}{\overline{\gamma}^{v}\Gamma(v)} e^{-\frac{\gamma_{\mathrm{v}}}{\overline{\gamma}}}, \quad \gamma_{\mathrm{v}} \geq 0, \qquad f(\gamma_{\mathrm{s}}) = \frac{\gamma_{\mathrm{s}}^{s-1}}{\overline{\gamma}^{s}\Gamma(s)} e^{-\frac{\gamma_{\mathrm{s}}}{\overline{\gamma}}}, \quad \gamma_{\mathrm{s}} \geq 0 \tag{28}$$

where $\overline{\gamma} = \mathbb{E}(\frac{|HS|^2}{\delta^2})$ denotes average SNR over multiple CRs, and $f(\cdot)$ denotes a generic probability density function (PDF) of its argument.

The average probabilities of false alarm and detection for MS^3 are often solved by averaging $P_{\mathrm{f},k}$ in (25) and $P_{\mathrm{d},k}$ in (26) over all possible SNRs, respectively.

Theorem 2: If the magnitudes of received signals at different CRs follow Rayleigh distribution, the average probabilities of false alarm ($\overline{P_{\mathrm{f},k}}$) and detection ($\overline{P_{\mathrm{d},k}}$) in MS^3 will have the following bounds

$$\frac{\Gamma(Jv, \frac{\lambda_k}{2})}{\Gamma(Jv)} \leq \overline{P_{\mathrm{f},k}} \leq \Theta(s, Jv, \psi, \overline{\gamma}[k], \lambda_k) \tag{29}$$

$$\overline{P_{\mathrm{d},k}} \geq \Theta(v, Jv, \psi, \overline{\gamma}[k], \lambda_k) \tag{30}$$

where $\Theta(x, Jv, \psi, \overline{\gamma}, \lambda)$ is defined as

$$\Theta = \left(1 + \frac{\psi\overline{\gamma}}{2}\right)^{-x} \sum_{n=0}^{\infty} \mathrm{C}_{n+x-1}^{n} \left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n} \frac{\Gamma\left(n + Jv, \frac{\lambda}{2}\right)}{\Gamma\left(n + Jv\right)} \tag{31}$$

in which C_a^b denotes the binomial coefficient, i.e., $\mathrm{C}_a^b = \frac{b!}{a!(b-a)!}$.

The proof of Theorem 2 is given in Appendix D.

Remark 2: From Theorem 2, we can see that $0 \leq \Theta \leq 1$, because the term $\frac{\Gamma(a,b)}{\Gamma(a)} \in [0, 1]$ and the remaining terms can be simplified to 1. In addition, it can be proved that Θ is a monotonically increasing function with respect to ψ, $\overline{\gamma}$, and x. Therefore, both probabilities will either increase or remain the same when the average sampling rate and the average SNR increase, more sampling channels will lead to a higher probability of detection, and the average probability of false alarm can be reduced with smaller s.

Remark 3: Because (31) contains infinite sums, its computational complexity is directly related to the number of computed terms that are required in order to obtain a specific accuracy. As the

number of computed terms, i.e., P, varies, the truncation error can be written as

$$T_{\Theta}(P) = \left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-x}\sum_{n=P}^{\infty}\mathrm{C}_{n+x-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n}\frac{\Gamma\left(n+Jv,\frac{\lambda}{2}\right)}{\Gamma\left(n+Jv\right)} \quad (32)$$

$$\leq \left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-x}\sum_{n=P}^{\infty}\mathrm{C}_{n+x-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n} \quad (33)$$

$$= 1-\left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-x}\sum_{n=0}^{P-1}\mathrm{C}_{n+x-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n} \quad (34)$$

where the inequality of (33) holds because $\frac{\Gamma(n,\frac{\lambda}{2})}{\Gamma(n)} \leq 1$, and (34) is obtained by using the binomial expansion. It can be shown that (31) converges very quickly. For example, in order to achieve double-precision accuracy, only $P = 30 \sim 40$ calculated terms are required; therefore the bounds are tractable. To solve for the detection threshold λ_k, we could use the lower bound on $P_{\mathrm{f},k}$ in (30). This is because the lower bound can approximate $P_{\mathrm{f},k}$ very well as analyzed in Appendix D and also verified by Fig. 3.

B. Log-normal Distribution

The strength of the transmitted primary signal is also affected by shadowing from buildings, hills, and other objects. A common model is that the received power fluctuates with a log-normal distribution. In such a scenario, the PDF of the SNR at CR i, i.e., $f(\gamma_i)$, is given by

$$f(\gamma_i) = \frac{\xi}{\sqrt{2\pi}\sigma_i\gamma_i}\exp\left(-\frac{(10\log_{10}(\gamma_i)-\overline{\gamma}_i)^2}{2\sigma_i^2}\right), \quad \gamma_i > 0 \quad (35)$$

where $\xi = 10/\ln(10)$, and σ_i (dB) denotes the standard deviation of $10\log_{10}\gamma_i$ at CR i. Note that the PDF in (35) can be closely approximated by a Wald distribution [27], [32]:

$$f(\gamma_i) = \sqrt{\frac{\eta_i}{2\pi}}\gamma_i^{-3/2}\exp\left(-\frac{\eta_i(\gamma_i-\theta_i)^2}{2\theta_i^2\gamma_i}\right), \quad \gamma_i > 0 \quad (36)$$

where $\theta_i = \mathbb{E}(\gamma_i)$ denotes the expectation of γ_i, and η_i is the shape parameter for CR i. Via the method of moments, the parameters η_i, θ_i and $\overline{\gamma}_i$, σ_i are related as follows:

$$\theta_i = \exp\left(\frac{\overline{\gamma}_i}{\xi}+\frac{\sigma_i^2}{2\xi^2}\right), \qquad \eta_i = \frac{\theta_i}{\exp(\frac{\sigma_i^2}{\xi^2})-1}. \quad (37)$$

In the proposed system, the condition $\frac{\eta_i}{\theta_i^2} = \frac{\mathbb{E}(\gamma_i)}{\mathrm{Var}(\gamma_i)} = b$ (constant) can be satisfied. Thus, γ_{s}

and $\gamma_{\rm v}$ will also follow the Wald distribution [33]. The PDFs of $\gamma_{\rm s}$ and $\gamma_{\rm v}$ are given by

$$f(\gamma_{\rm s}) = \sqrt{\frac{s\eta}{2\pi}}\gamma_{\rm s}^{-3/2}\exp\left(-\frac{\eta(\gamma_{\rm s}-s\theta)^2}{2s\theta^2\gamma_{\rm s}}\right), \quad \gamma_{\rm s}>0 \tag{38}$$

$$f(\gamma_{\rm v}) = \sqrt{\frac{v\eta}{2\pi}}\gamma_{\rm v}^{-3/2}\exp\left(-\frac{\eta(\gamma_{\rm v}-v\theta)^2}{2v\theta^2\gamma_{\rm v}}\right), \quad \gamma_{\rm v}>0 \tag{39}$$

where η and θ denote the averages of η_i and θ_i, respectively.

Theorem 3: If the magnitudes of received signals at different CRs follow log-normal distribution, the average probabilities of false alarm ($\widetilde{P_{{\rm f},k}}$) and detection ($\widetilde{P_{{\rm d},k}}$) in MS[3] will be bounded as

$$\frac{\Gamma(Jv,\frac{\lambda_k}{2})}{\Gamma(Jv)} \le \widetilde{P_{{\rm f},k}} \le \Lambda(s, Jv, \psi, \lambda_k, \theta[k], \eta[k]) \tag{40}$$

$$\widetilde{P_{{\rm d},k}} \ge \Lambda(v, Jv, \psi, \lambda_k, \theta[k], \eta[k]) \tag{41}$$

where $\Lambda(x, Jv, \psi, \lambda, \theta, \eta)$ is defined by

$$\Lambda = \sqrt{\frac{2x\eta}{\pi}}e^{\frac{\eta}{\theta}}\sum_{n=0}^{\infty}\frac{\left(\frac{\psi}{2}\right)^n\Gamma\left(n+Jv,\frac{\lambda}{2}\right)}{n!\Gamma\left(n+Jv\right)}\left(\sqrt{\frac{x^2\eta\theta^2}{x\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}}{\rm K}_{n-\frac{1}{2}}\left(\sqrt{\frac{\eta(x\psi\theta^2+\eta)}{\theta^2}}\right) \tag{42}$$

in which ${\rm K}_{n-\frac{1}{2}}(a)$ denotes the modified Bessel function of the second kind with order $n-\frac{1}{2}$.

The proof of Theorem 3 is given in Appendix E.

Remark 4: Because (42) contains infinite sums, the truncation error $T_\Lambda(P)$ must be considered. Similar to (33), the truncation error can be written as

$$\begin{aligned} T_\Lambda(P) &\le \sqrt{\frac{2x\eta}{\pi}}e^{\frac{\eta}{\theta}}\sum_{n=P}^{\infty}\frac{\left(\frac{\psi}{2}\right)^n\left(\sqrt{\frac{x^2\eta\theta^2}{x\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}}}{n!}{\rm K}_{n-\frac{1}{2}}\left(\frac{\sqrt{\eta(x\psi\theta^2+\eta)}}{\theta^2}\right) \\ &= 1-\sqrt{\frac{2x\eta}{\pi}}e^{\frac{\eta}{\theta}}\sum_{n=0}^{P-1}\frac{\left(\frac{\psi}{2}\right)^n\left(\sqrt{\frac{x^2\eta\theta^2}{x\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}}}{n!}{\rm K}_{n-\frac{1}{2}}\left(\frac{\sqrt{\eta(x\psi\theta^2+\eta)}}{\theta^2}\right). \end{aligned} \tag{43}$$

It can be shown that (43) decreases to zero very quickly as P increases, therefore the bounds in Theorem 3 are easy to compute.

V. Simulation results

In our simulations, we assume that the CRs are organized as shown in Fig. 1 and adopt the following configurations unless otherwise stated. We use the wideband analog signal model in

[34] and thus the received signal $x_{\mathrm{c},i}(t)$ at CR i has the form:

$$x_{\mathrm{c},i}(t) = \sum_{l=1}^{N_b} |H_{i,l}| \sqrt{E_l} B_l \cdot \mathrm{sinc}\left(B_l(t-\Delta)\right) \cdot \cos\left(2\pi f_l(t-\Delta)\right) + z(t) \tag{44}$$

where $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$, Δ denotes a random time offset, $z(t)$ is AWGN, i.e., $z(t) \sim \mathcal{N}(0,1)$, E_l is the transmit power at PU, and $H_{i,l}$ denotes the discrete frequency response between the PU and CR i in subband l. We generate $v = 22$ independent channels according to the fading environment, and regenerate them for next observation time. The received signal $x_{\mathrm{c},i}(t)$ consists of $N_b = 6$ non-overlapping subbands. The l-th subband is in the frequency range of [$f_l - \frac{B_l}{2}$, $f_l + \frac{B_l}{2}$], where the bandwidth $B_l = 1 \sim 10$ MHz and f_l denotes the center frequency. The center frequency of the subband l is randomly located within $[\frac{B_l}{2}, W - \frac{B_l}{2}]$ (i.e., $f_l \in [\frac{B_l}{2}, W - \frac{B_l}{2}]$), where the overall signal bandwidth $W = 10$ GHz. If the wideband signal were sampled at the Nyquist rate $f = 2W$ for $T = 20$ μs, after segment division with $J = 5$, the number of Nyquist samples per segment would be $N = 80,000$; thus, using FFT-based approach, the frequency resolution is $\frac{1}{T/J} = 0.25$ MHz. In MS^3, the received signal is sampled by using different sub-Nyquist rates at different CRs. To be specific, the numbers of samples in multiple CRs are chosen by using Theorem 1 and we choose the first prime $M_1 \approx a\sqrt{N}$ ($a \geq 1$) and its $v-1$ neighboring and consecutive primes. The spectral observations are obtained by applying an FFT to these sub-Nyquist samples in each channel. **Then the signal energy is calculated in the spectral domain using (16), and the energy vectors are transmitted from the CRs to the FC using dedicated common control channels. These channels are assumed to be Rayleigh block fading channels (constant channel gains over one time block) corrupted by circularly symmetric complex Gaussian noise with zero mean and unit variance. In addition, we consider that the channel power gain of the common control channel is normalized to unit and the average SNR as received at the FC is 15 dB.** In the FC, we form the test statistic by using (18). We define the compression rate as the ratio between the number of samples at the sub-Nyquist rate and the number of samples at the Nyquist rate, i.e., $\frac{M}{N}$ where M denotes the average number of sub-Nyquist samples at CRs. Spectrum sensing results are obtained by using the decision rule (19) and varying the detection threshold λ_k.

In Fig. 3, we verify the theoretical results in (25)-(26), (29)-(30), and (40)-(41) by comparing them with the simulated results. It shows that the lower bound on the probability of false alarm can tightly predict the simulated results while the upper bound seems relatively loose. It is because that the assumption (i.e., all s components in the Nyquist DFT spectrum will be mapped to the same location when the signal is sub-Nyquist sampled) for deriving the upper bound can rarely occur. Fig. 3 also illustrates that the lower bound on the probability of detection can successfully predict the trend of simulated results. Comparing the faded signal cases with the non-faded signal case, it is found that, when combining faded signals, the probability of detection declines more slow than that combining non-faded signals. This is more obvious for the case of combining signals following log-normal distribution as shown in Fig. 3(c).

Fig. 4 shows the receiver operating characteristic (ROC) curves of MS^3 when combining non-faded and faded signals. When the average SNR as received at CRs is 5 dB, the performance of MS^3 combining faded signals is roughly the same as that of combining non-faded signals. This is because the strength of the signal is mostly masked by the noise. In contrast, the detection performance of MS^3 combining non-faded signals outperforms that of combining faded signals when SNR=10 dB. In addition, it is seen that the performance of MS^3 combining log-normal shadowed signals is the poorest. Nonetheless, even for log-normal shadowed signals, MS^3 has a probability of nearly 90% for detecting the presence of PUs when the probability of false alarm is 10%, with the compression rate of $\frac{M}{N} = 0.0219$. To investigate the influence of s and SNR, we use Fig. 5 to show the performance of MS^3 when the received signals are faded according to Rayleigh distribution with different values of s (proportional to the number of subbands). We see that, as the number of subbands decreases, the detection performance improves for the same SNR. The performance improvement of MS^3 stems from that, for a fixed number of sampling channels, decreasing s makes it easier to distinguish the occupied frequencies from the aliased frequencies as discussions in Section III-C.

Fig. 6(a) depicts the influence of the standard deviation when the MS^3 system combines log-normal shadowed signals. It can be seen that a larger standard deviation will lead to worse detection performance for the MS^3 system. It is because a larger σ is equivalent to a longer tail in

the log-normal distribution, thus making the detection more difficult. **In Fig. 6(b), we compare the performance of MS^3 with that of Nyquist systems. In the Nyquist system type I, each CR is given an orthogonal subband (wideband spectrum is divided into several equal-length subbands) to sense using Nyquist rate, while their decisions are sent back to the FC. In the Nyquist system type II, we assume that each CR must sense all wideband spectrum non-cooperatively, thus requiring multiple standard ADCs in each node to cover all wideband spectrum. After signal sampling, all measurements are sent back to the FC, where equal gain combining approach is adopted to fuse data and then energy detection is used for spectrum sensing. Fig. 6(b) shows that the proposed system has superior performance to the Nyquist system type I, but inferior performance to the Nyquist system type II. The poor performance of the Nyquist system type I mainly results from the lack of spatial diversity gain. In the Nyquist system type I, each subband is only sensed by one CR as each CR is given an orthogonal subband to sense, which cannot take advantage of spatial diversity. In contrast, both the proposed system and the Nyquist system type II are monitoring each subband using several CRs, thus taking advantage of spatial diversity. It can also be seen that the Nyquist system type II has marginal performance gain over the proposed system, however, at the expense of much higher implementation complexity as discussed below.**

In Table I, we compare the implementation complexity of MS^3 with that of the Nyquist systems, when the received signals at different CRs are faded according to Rayleigh distribution. Here, we consider the comparison metric: the number of same-sampling-rate ADCs for achieving $P_{\mathrm{d}} \geq 90\%$ and $P_{\mathrm{f}} \leq 10\%$, because practical CRs often have requirements on the probabilities of detection and false alarm to secure the performance of both CRs and PUs. We can see that, when there exist 10 CRs, MS^3 requires each CR equipped with a single ADC with an average sampling rate of 957.54 MHz; thus, the whole CR network only requires 10 low-rate ADCs. In contrast, the Nyquist system type I requires 21 ADCs in total, because of 21×957.54 MHz≈ 20 GHz for covering 10 GHz spectrum based on Nyquist sampling theorem. In the Nyquist system type II, 210 ADCs (with the average sampling rate 957.54 MHz) will be required because each CR will require 21 ADCs similar to the Nyquist system type I. **Thus, the system complexity**

of MS^3 is approximately half of that of the Nyquist system type I and much less than that of the Nyquist system type II.

In Fig. 7, we choose the CS-based system in [13] as a benchmark system due to its high impact and outstanding performance. The comparison between the proposed MS^3 system and the benchmark system is provided. We assume that $v = 22$ CRs are collaborating for wideband spectrum sensing in both systems, in order to increase the reliability of spectrum sensing by exploiting spatial diversity. We can see from Fig. 7(a) that MS^3 outperforms the CS-based system for every compression rate. **In Fig. 7(b), it is seen that, compared with the benchmark system, MS^3 has better compression capability. Using MS^3, the probability of successful sensing becomes larger than 90% when the compression rate $\frac{M}{N} \geq 0.023$. In contrast, the benchmark system can achieve the probability of successful sensing 90% only when the compression rate $\frac{M}{N} \geq 0.045$.** Furthermore, as shown in Table II, we can find that the computational complexity of MS^3 is $\mathcal{O}(N \log N)$ due to the energy detection with FFT operations, rather than $\mathcal{O}\left(N(M + \log N)\right)$ in the CS-based system, where M is usually much larger than $\log N$. The complexity of the CS-based system is caused by both the matrix multiplication operations and the FFT operations for spectral recovery. To sum up, with the same computational resources, MS^3 has a relatively smaller spectrum sensing overhead than the CS-based system, not only because of the better compression capability (less data transmission results in shorter transmission time), but also due to the lower computational complexity.

VI. Conclusions

In this paper, we have presented a novel system, i.e., MS^3, for wideband spectrum sensing in CR networks. MS^3 can relax the wideband spectrum sensing requirements of CRs due to its capability of sub-Nyquist sampling. It has been shown that, using sub-Nyquist samples, the wideband spectrum can be sensed in a collaborative manner without spectral recovery, leading to a high energy-efficiency and a low spectrum sensing overhead. Moreover, we have derived closed-form bounds for the performance of MS^3 when combining faded or shadowed signals.

Simulation results have verified the derived bounds on the probabilities of false alarm and

detection. It has also been shown that using partial measurements, MS^3 has superior performance even under low SNR scenarios. The performance of MS^3 improves as either the number of CRs or the average sampling rate increases. Compared to the existing wideband spectrum sensing methods, MS^3 not only provides computation and memory savings, but also reduces the hardware acquisition requirements and the energy costs at CRs.

APPENDIX A
RELATIONSHIP BETWEEN NYQUIST DFT SPECTRUM AND SUB-NYQUIST DFT SPECTRUM

Using the Poisson summation formula [29], (9), and (10), we obtain:

$$f_i \sum_{l\in\mathbb{Z}} X_{\mathrm{c},i}(w+f_i l) = \sum_{n\in\mathbb{Z}} y_i[n] e^{-\jmath 2\pi w n} = \sum_{n=0}^{M_i-1} y_i[n] e^{-\jmath 2\pi w n} = Y_i(w) \tag{45}$$

where $X_{\mathrm{c},i}(w) = \int_{-\infty}^{\infty} x_{\mathrm{c},i}(t) e^{-\jmath 2\pi w t} dt$. Similar to (45), by using (1) and (2), we can obtain:

$$f \sum_{l\in\mathbb{Z}} X_{\mathrm{c},i}(w+fl) = \sum_{n\in\mathbb{Z}} x_i[n] e^{-\jmath 2\pi w n} = \sum_{n=0}^{N-1} x_i[n] e^{-\jmath 2\pi w n} = X_i(w). \tag{46}$$

As the received signal is bandlimited and $f \geq 2W$, $X_i(w) = f X_{\mathrm{c},i}(w)$ holds for $w \in [-\frac{W}{2}, \frac{W}{2}]$. Substituting it to (45), we obtain $Y_i(w) = \frac{f_i}{f} \sum_{l=-\infty}^{\infty} X_i(w+f_i l)$. In a discrete form, we end up with:

$$Y_i[m] = \frac{M_i}{N} \sum_{l=-\infty}^{\infty} X_i[m+lM_i], \quad m = 0, 1, \cdots, M_i - 1. \tag{47}$$

APPENDIX B
PROBABILITY OF SIGNAL OVERLAP AT SUB-NYQUIST SAMPLING

As s spectral components are distributed over the frequency bins of $0, 1, \cdots, N-1$, the probability of the frequency bin k belonging to the spectral support Ω is $P = \Pr(k \in \Omega) = \frac{s}{N}$. Let q denote the number of spectral components overlapped on the frequency bin m, using (11) the probability of no signal overlap is given by

$$\Pr(q<2) = \Pr(q=0) + \Pr(q=1) = (1-P)^{\lceil \frac{N}{M_i} \rceil} + \binom{\lceil \frac{N}{M_i} \rceil}{1} P (1-P)^{\lceil \frac{N}{M_i} \rceil - 1} \tag{48}$$

where $\lceil \frac{N}{M_i} \rceil$ denotes the number of summations in (11). Substituting $P = \frac{s}{N}$ into (48) while choosing sub-Nyquist sampling rate in MS^3 such that $M_i = \sqrt{N}$, we obtain

$$\Pr(q < 2) = \left(\frac{N-s}{N}\right)^{\frac{N}{M_i}} + \frac{s}{M_i}\left(\frac{N-s}{N}\right)^{\frac{N-M_i}{M_i}} = \frac{(\frac{N-s}{N})^{\sqrt{N}}(N-s+s\sqrt{N})}{N-s}. \tag{49}$$

It can be tested that $\Pr(q < 2)$ approaches to 1 when choosing N such that $N \gg s$. Thus the probability of signal overlap approaches to zero under the condition we choose.

APPENDIX C
PROOF OF LEMMA 1

Let M_i and M_j denote the number of samples at CRs i and j, respectively. Using (12) and (20), we can represent the aliased frequencies projected from $k_1 \in \Omega$ by

$$g_i = |k_1|_{\bmod (M_i)} + lM_i = k_1 - hM_i + lM_i,\ h \neq l \tag{50}$$

$$g_j = |k_1|_{\bmod (M_j)} + \check{l}M_j = k_1 - \check{h}M_j + \check{l}M_j,\ \check{h} \neq \check{l} \tag{51}$$

where integers h and $\check{h}$ are quotients from modulo operations, and $l-h \in [-\lceil \frac{N}{M_i} \rceil + 1, \lceil \frac{N}{M_i} \rceil - 1]$, $\check{l} - \check{h} \in [-\lceil \frac{N}{M_j} \rceil + 1, \lceil \frac{N}{M_j} \rceil - 1]$, in which $\lceil \frac{N}{M_i} \rceil$ gives the smallest integer not less than $\frac{N}{M_i}$.

Avoiding $g_i = g_j$ is equivalent to avoiding $(l-h)M_i = (\check{l} - \check{h})M_j$. If M_i and M_j are different primes, the condition $\max(|l-h|) < M_j$ (i.e., $\lceil \frac{N}{M_i} \rceil - 1 < M_j$) will satisfy this. After simplification, the condition $M_iM_j > N$ is obtained. Moreover, if it holds for any two CRs, the case for more than two CRs will also hold.

APPENDIX D
PROOF OF THEOREM 2

If the received signals at CRs are Rayleigh faded, the lower bound on the average probability of false alarm will remain as it is independent of the SNR. Using (25), (27), and (28), the upper bound on the average probability of false alarm can be calculated by

$$\overline{P_{\mathrm{f},k}}^{\mathrm{up}} = \int_0^\infty Q_{Jv}\Big(\sqrt{\psi\gamma_{\mathrm{s}}}, \sqrt{\lambda_k}\Big) \frac{\gamma_{\mathrm{s}}^{s-1}}{\overline{\gamma}^s \Gamma(s)} e^{-\frac{\gamma_{\mathrm{s}}}{\overline{\gamma}}} d\gamma_{\mathrm{s}}. \tag{52}$$

Rewriting the Marcum Q-function by using (4.74) in [35] and (8.352-2) in [36], we obtain:

$$Q_{Jv}\left(\sqrt{\psi\gamma_{\rm s}},\sqrt{\lambda_k}\right)=\sum_{n=0}^{\infty}\frac{\left(\frac{\psi\gamma_{\rm s}}{2}\right)^n e^{-\frac{\psi\gamma_{\rm s}}{2}}}{n!}\frac{\Gamma(n+Jv,\frac{\lambda_k}{2})}{\Gamma(n+Jv)}. \tag{53}$$

Substituting (53) into (52), we can rewrite (52) as

$$\overline{P_{{\rm f},k}}^{\rm up}=\frac{1}{\overline{\gamma}^s}\sum_{n=0}^{\infty}\frac{\left(\frac{\psi}{2}\right)^n\Gamma(n+Jv,\frac{\lambda_k}{2})}{n!(s-1)!\Gamma(n+Jv)}\int_0^{\infty}\gamma_{\rm s}^{n+s-1}e^{-\frac{\psi\gamma_{\rm s}}{2}-\frac{\gamma_{\rm s}}{\overline{\gamma}}}d\gamma_{\rm s}. \tag{54}$$

Calculating the integral by using (3.351-3) in [36], we end up with

$$\overline{P_{{\rm f},k}}^{\rm up}=\left(1+\frac{\psi\overline{\gamma}}{2}\right)^{-s}\sum_{n=0}^{\infty}{\rm C}_{n+s-1}^{n}\left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^n\frac{\Gamma\left(n+Jv,\frac{\lambda_k}{2}\right)}{\Gamma\left(n+Jv\right)}. \tag{55}$$

Similarly, we can obtain the lower bound on the average probability of detection.

APPENDIX E
PROOF OF THEOREM 3

If the received signals are shadowed according to log-normal distribution, the lower bound on $\widetilde{P_{{\rm f},k}}$ in (40) will remain. By (38), the upper bound on the probability of false alarm can be given by

$$\widetilde{P_{{\rm f},k}}^{u}=\int_0^{\infty}Q_{Jv}\Big(\sqrt{\psi\gamma_{\rm s}},\sqrt{\lambda_k}\Big)\sqrt{\frac{s\eta}{2\pi}}\gamma_{\rm s}^{-3/2}\exp\left(-\frac{\eta(\gamma_{\rm s}-s\theta)^2}{2s\theta^2\gamma_{\rm s}}\right)d\gamma_{\rm s}. \tag{56}$$

Substituting (53) into (56), we calculate $\widetilde{P_{{\rm f},k}}^{u}$ as

$$\widetilde{P_{{\rm f},k}}^{u}=\sqrt{\frac{s\eta}{2\pi}}\sum_{n=0}^{\infty}\frac{\left(\frac{\psi}{2}\right)^n\Gamma\left(n+Jv,\frac{\lambda_k}{2}\right)}{n!\Gamma\left(n+Jv\right)}\int_0^{\infty}\gamma_{\rm s}^{n-\frac{3}{2}}\exp\left(-\frac{s\psi\theta^2+\eta}{2s\theta^2}\gamma_{\rm s}-\frac{s\eta}{2\gamma_{\rm s}}+\frac{\eta}{\theta}\right)d\gamma_{\rm s}. \tag{57}$$

Using (3.471-9) in [36] for calculating the integral in (57), we obtain

$$\widetilde{P_{{\rm f},k}}^{u}=\sqrt{\frac{2s\eta}{\pi}}e^{\frac{\eta}{\theta}}\sum_{n=0}^{\infty}\frac{\left(\frac{\psi}{2}\right)^n\Gamma\left(n+Jv,\frac{\lambda_k}{2}\right)}{n!\Gamma\left(n+Jv\right)}\left(\sqrt{\frac{s^2\eta\theta^2}{s\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}}{\rm K}_{n-\frac{1}{2}}\left(\sqrt{\frac{\eta(s\psi\theta^2+\eta)}{\theta^2}}\right). \tag{58}$$

Likewise, the lower bound on the average probability of detection can be approximated.

REFERENCES

[1] M. A. McHenry, "NSF spectrum occupancy measurements project summary," Shared Spectrum Company, Tech. Rep., Aug. 2005.

[2] S. Stotas and A. Nallanathan, "On the outage capacity of sensing-enhanced spectrum sharing cognitive radio systems in fading channels," *IEEE Transactions on Communications*, vol. 59, no. 10, pp. 2871–2882, October 2011.

[3] Z. Chen, C.-X. Wang, X. Hong, J. Thompson, S. Vorobyov, X. Ge, H. Xiao, and F. Zhao, "Aggregate interference modeling in cognitive radio networks with power and contention control," *IEEE Transactions on Communications*, vol. 60, no. 2, pp. 456–468, Feb. 2012.

[4] H. Sun, D. I. Laurenson, J. S. Thompson, and C.-X. Wang, "A novel centralized network for sensing spectrum in cognitive radio," in *Proc. IEEE International Conference on Communications*, Beijing, China, May 2008, pp. 4186–4190.

[5] Q. Chen, M. Motani, W.-C. Wong, and A. Nallanathan, "Cooperative spectrum sensing strategies for cognitive radio mesh networks," *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 56–67, Feb. 2011.

[6] M.-H. Yoon, Y. Shin, H.-K. Ryu, and J.-M. Woo, "Ultra-wideband loop antenna," *Electronics Letters*, vol. 46, no. 18, pp. 1249–1251, Sept. 2010.

[7] Z.-C. Hao and J.-S. Hong, "Highly selective ultra wideband bandpass filters with quasi-elliptic function response," *IET Microwaves, Antennas Propagation*, vol. 5, no. 9, pp. 1103–1108, 2011.

[8] A. Bevilacqua and A. M. Niknejad, "An ultrawideband CMOS low-noise amplifier for 3.1-10.6-GHz wireless receivers," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2259–2268, Dec. 2004.

[9] [Online]. Available: www.national.com/pf/DC/ADC12D1800.html

[10] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.

[11] ——, "Wideband spectrum sensing in cognitive radio networks," in *Proc. IEEE International Conference on Communications*, Beijing, China, May 2008, pp. 901–906.

[12] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proc. IEEE Cognitive Radio Oriented Wireless Networks and Communications*, Mykonos Island, Greece, June 2006, pp. 1–5.

[13] ——, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hawaii, USA, April 2007, pp. 1357–1360.

[14] Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-nyquist sampling for wideband spectrum sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 1, pp. 58–69, Feb. 2012.

[15] F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 37–48, Feb. 2011.

[16] Y. Wang, Z. Tian, and C. Feng, "Sparsity order estimation and its application in compressive spectrum sensing for cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2116–2125, June 2012.

[17] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Taipei, April 2009, pp. 2337–2340.

[18] H. Sun, A. Nallanathan, J. Jiang, D. I. Laurenson, C.-X. Wang, and H. V. Poor, "A novel wideband spectrum sensing system for distributed cognitive radio networks," in *Proc. IEEE Global Telecommunications Conference*, Houston, TX, USA, Dec. 2011, pp. 1–6.

[19] Y. Wang, A. Pandharipande, and G. Leus, "Compressive sampling based MVDR spectrum sensing," in *2010 2nd International Workshop on Cognitive Information Processing (CIP)*, June 2010, pp. 333–337.

[20] V. Havary-Nassab, S. Hassan, and S. Valaee, "Compressive detection for wide-band spectrum sensing," in *in Proc. of IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP 2010)*, Dallas, TX, USA, Mar. 2010, pp. 3094–3097.

[21] U. Nakarmi and N. Rahnavard, "Joint wideband spectrum sensing in frequency overlapping cognitive radio networks using distributed compressive sensing," in *Proc. Military Communications Conference*, Baltimore, MD, USA, Nov. 2011, pp. 1035–1040.

[22] Y. Zhang, Y. Liu, and Q. Wan, "Fast compressive wideband spectrum sensing based on dictionary linear combination," in *Proc. IEEE WiCOM*, Wuhan, China, Sept. 2011, pp. 1–4.

[23] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4775–4789, sept. 2012.

[24] Y. Liu and Q. Wan, "Enhanced compressive wideband frequency spectrum sensing for dynamic spectrum access," *EURASIP Journal on Advances in Signal Processing*, vol. 2012(177), 2012.

[25] H. Sun, A. Nallanathan, and J. Jiang, "Adaptive compressive spectrum sensing in distributed cognitive radio networks," in *Proc. IEEE Asia Pacific Wireless Communication Symposium*, Singapore, August 2011, pp. 1–5.

[26] H. Sun, A. Nallanathan, J. Jiang, and H. V. Poor, "Compressive autonomous sensing (CASe) for wideband spectrum sensing," in *Proc. of IEEE International Conference on Communications*, Ottawa, Canada, June 2012, pp. 5953–5957.

[27] H. Sun, D. Laurenson, and C.-X. Wang, "Computationally tractable model of energy detection performance over slow fading channels," *IEEE Communications Letters*, vol. 14, no. 10, pp. 924–926, Oct. 2010.

[28] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 118–129, 2008.

[29] M. A. Pinsky, *Introduction to Fourier Analysis and Wavelets*. Providence, Rhode Island, USA: American Mathematical Society, 2002.

[30] P. Cantrell and A. Ojha, "Comparison of generalized Q-function algorithms," *IEEE Transactions on Information Theory*, vol. 33, no. 4, pp. 591–596, July 1987.

[31] C. W. Helstrom, "Approximate inversion of Marcum's Q-function," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 317–319, Jan. 1998.

[32] A. H. Marcus, "Power sum distributions: an easier approach using the Wald distribution," *Journal of American Statistical Association*, vol. 71, pp. 237–238, 1976.

[33] R. S. Chhikara and J. L. Folks, *The Inverse Gaussian Distribution: Theory, Methodology, and Applications*. New York: Marcel Dekker Inc., 1989.

[34] M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375 –391, april 2010.

[35] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. New York: John Wiley & Sons, Inc., Dec. 2004.

[36] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed., A. Jeffrey, Ed. New York: Academic Press, Inc., 1994.

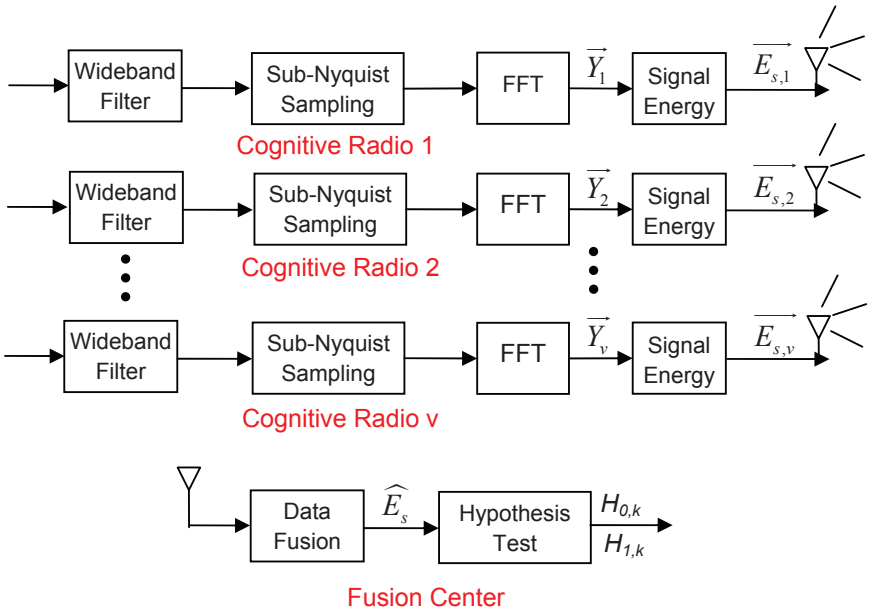

Fig. 1. Block diagram of multi-rate sub-Nyquist spectrum sensing (MS^3) system.

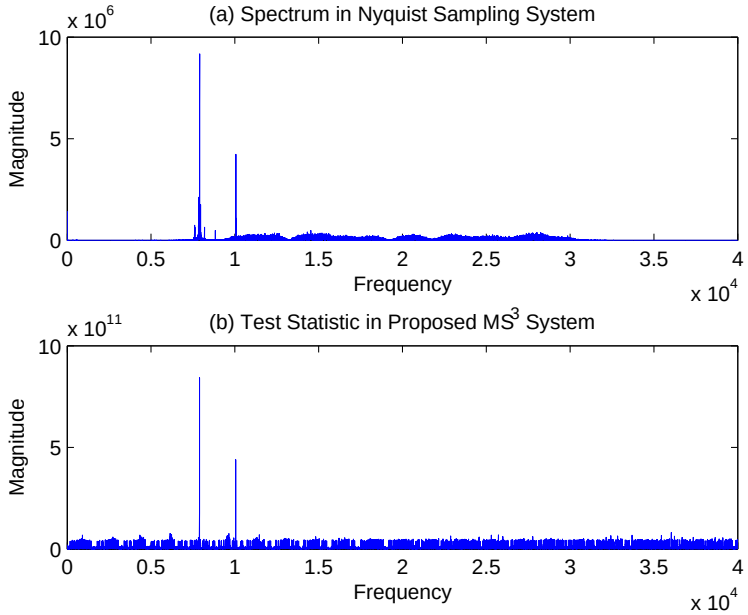

Fig. 2. Illustration of the proposed system for sensing practical ASTC DTV signal: (a) Spectrum of an ASTC DTV signal in Nyquist sampling system, and (b) Test statistic in the proposed system. Here, we use ASTC DTV signal WAS_3_27_06022000_REF and assume the number of CRs $v = 22$. The number of Nyquist samples is $N = 80,000$, and the numbers of sub-Nyquist samples in MS^3 system are consecutive primes $M_1 = 1613, M_2 = 1619, \cdots, M_{22} = 1783$. The average compression rate is calculated to $\frac{M}{N} = 2.12\%$.

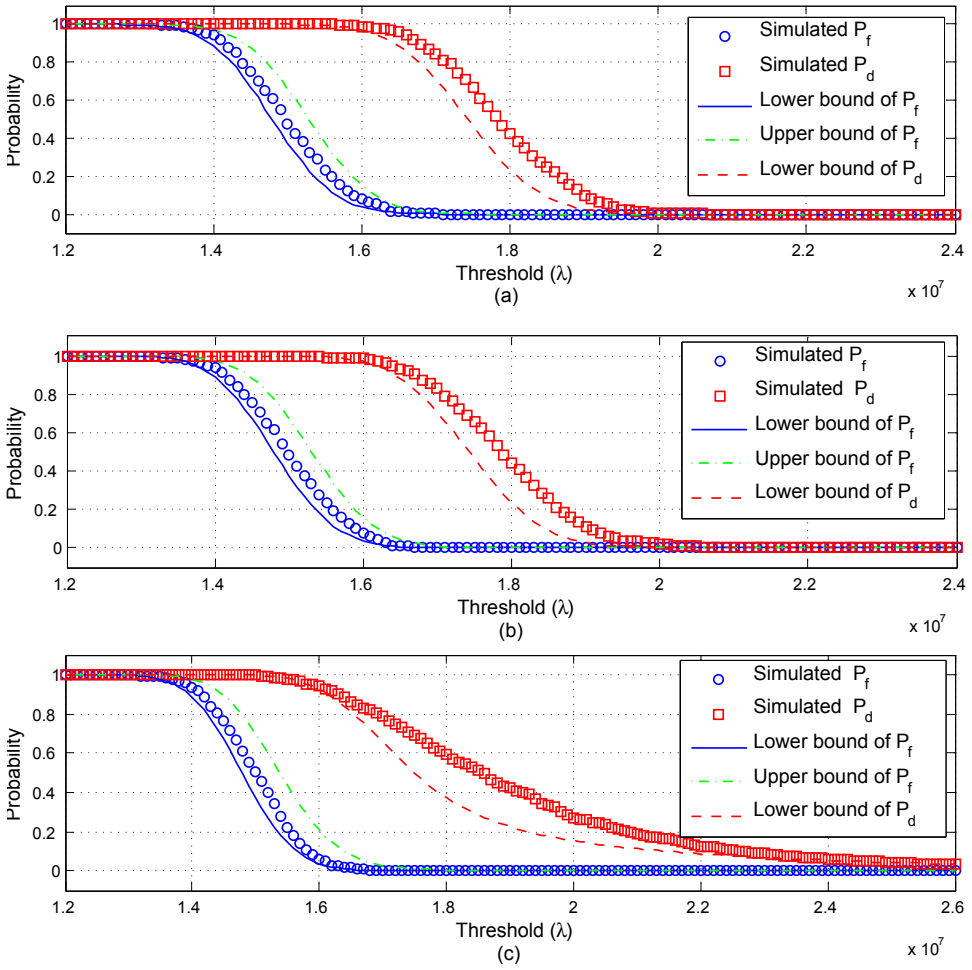

Fig. 3. Comparisons of simulation results and analytical results for the probabilities of false alarm and detection when MS[3] combining (a)non-faded signals, (b) Rayleigh faded signals, and (c) Log-normal shadowed signals with the received SNR$= 5$ dB (at CRs) and $\sigma = 4$ dB.

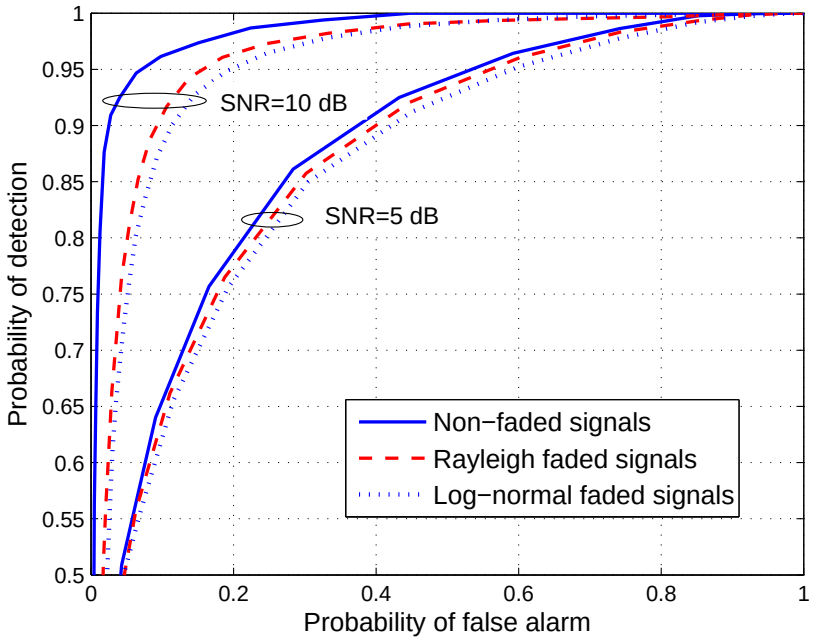

Fig. 4. Receiver operating characteristic curves of MS^3 for combining non-faded signals or faded signals when the compression rate $\frac{M}{N} = 0.0219$ and the number of segments $J = 5$. The wideband signal is observed by 22 CRs at different sampling rates (the average sampling rate is 448.68 MHz).

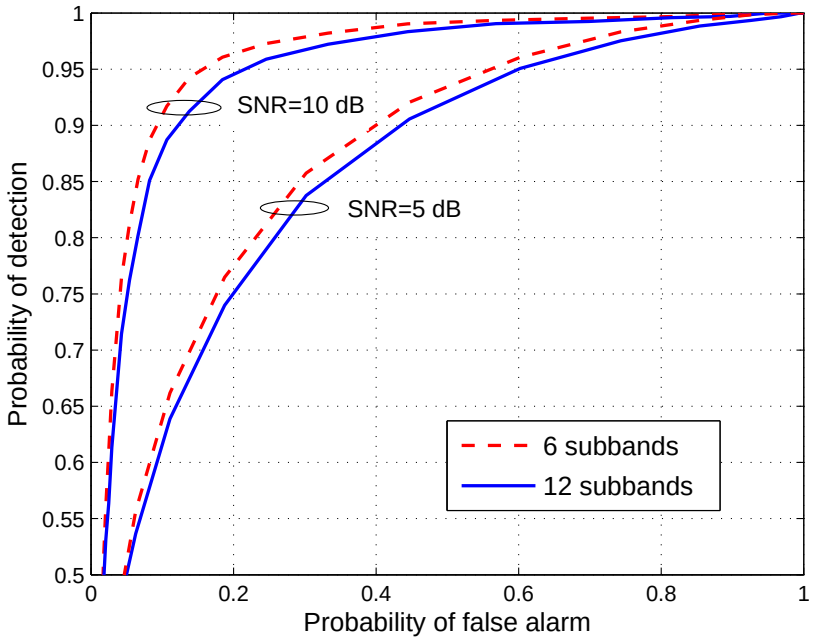

Fig. 5. The performance of MS^3 for combining Rayleigh faded signals with $v = 22$ and $\frac{M}{N} = 0.0219$, when the received SNR at CRs and the number of subbands change.

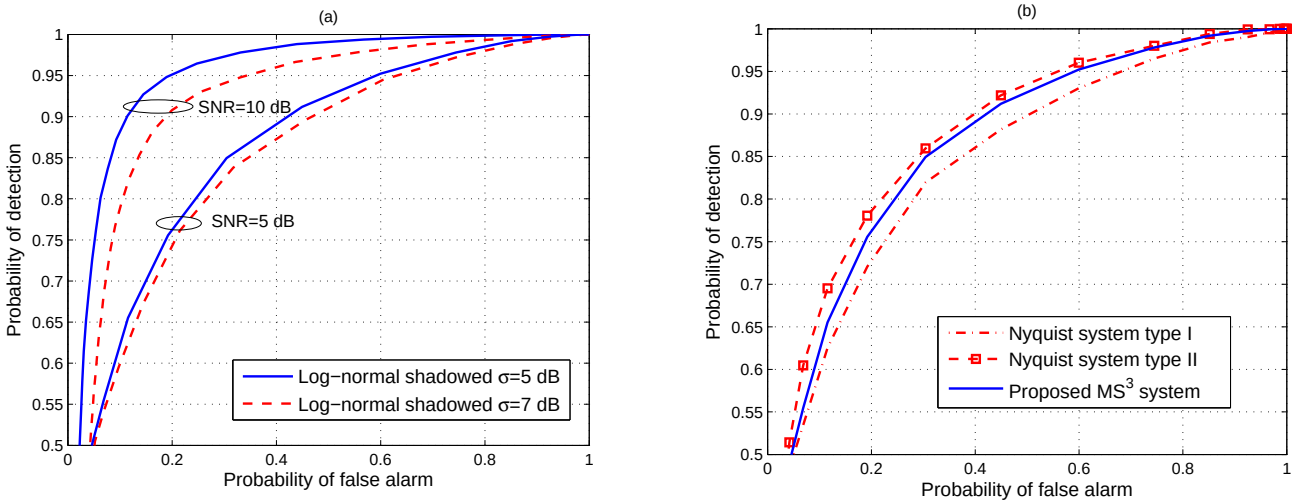

Fig. 6. Receiver operating characteristic curves for combining log-normal faded signals: (a) MS^3 when the standard deviation, i.e., σ, and the average SNR at CRs vary, (b) comparison between MS^3 and Nyquist systems when $\sigma = 5$ dB and the average SNR as received at CRs is 5 dB. In simulations, the wideband signal is sampled at different sampling rates by 22 ADCs with the average sampling rate of 448.68 MHz, and the compression rate is $\frac{M}{N} = 0.0219$ and $J = 5$.

TABLE I
IMPLEMENTATION COMPLEXITY COMPARISON OF MS^3 AND THE NYQUIST SYSTEMS WHEN THE RECEIVED SIGNALS ARE FADED ACCORDING TO RAYLEIGH DISTRIBUTION WITH TEN DECIBEL RECEIVED SNR AT CRS.

Number of CRs for Wideband Spectrum Sensing (v)	10	20	30	40
Required Number of ADCs in the Proposed MS^3 System	10	20	30	40
Required Number of ADCs in the Nyquist System Type I	21	40	58	74
Required Number of ADCs in the Nyquist System Type II	21×10	40×20	58×30	74×40
Average Sampling Rate of ADCs in the Above Systems ($\overline{f}$ in MHz)	957.54	513.08	350.34	276.77
Sample Reduction Rate I ($\frac{\text{Total samples in MS}^3}{\text{Total samples in Nyquist type I}}$)	47.62%	50%	51.72%	54.05%
Sample Reduction Rate II ($\frac{\text{Total samples in MS}^3}{\text{Total samples in Nyquist type II}}$)	4.76%	2.5%	1.72%	1.35%

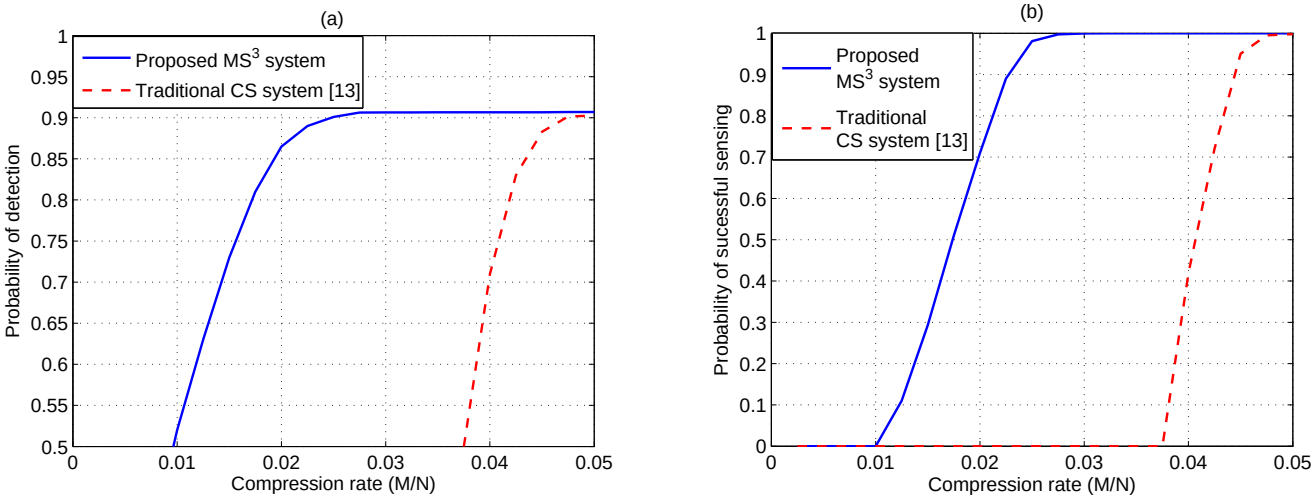

Fig. 7. Comparison between MS^3 and CS-based system [13]: (a) the probability of detection when the probability of false alarm is set to 10%, and (b) the probability of successful sensing which is defined as the probability of achieving both $P_{\rm d} \geq 90\%$ and $P_{\rm f} \leq 10\%$. In simulations, the average SNR as received at CRs is 10 dB and the number of CRs is $v = 22$.

TABLE II
COMPARISONS OF WIDEBAND SPECTRUM SENSING TECHNIQUES.

Approach	Compression Capability	ADC/DSP Type	Implementation Complexity	Computational Complexity
Wavelet detection	$\times$	Nyquist	low	$\mathcal{O}(N \log N)$
Multiband joint detection	$\times$	Nyquist	high	$\mathcal{O}\left(N \log N\right)$
CS-based detection	$\surd$	sub-Nyquist	medium	$\mathcal{O}\left(N(M + \log N)\right)$
Proposed system	$\surd$	sub-Nyquist	low	$\mathcal{O}(N \log N)$